Kaufman, Claire

From: Kaufman, Claire
Sent: Wednesday, August 13, 2003 10:50 AM
To: STIC-Biotech/ChemLib
Subject: sequence search 09/990,940

SEQUENCE SEARCH REQUEST

NAME: CLAIRE KAUFMAN AU: 1646 MAILBOX: 10D19

SERIAL NUMBER: **09/990,940** DATE: 8/13/03

Please search SEQ ID NO:17 and 18 and
oligo search both 17 and 18 (I need 50 contiguous amino acids or 100 contiguous nucleotides). Please search in commercial and interference database.

Please put results on disk.

Thanks!

Claire Kaufman
AU 1646, 305-5791

```
AC   Q8BHH0; PRELIMINARY;      PRT;   365 AA
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical rhodopsin-like GPCR superfamily containing protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion, and Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK048439; BAC33337.1; -.
DR   EMBL; AK051723; BAC34735.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   365 AA;  41759 MW;  1EB7E5369632ED56 CRC64;

  Query Match             88.2%;  Score 1905;  DB 11;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-160;
  Matches  365;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps 0

Qy         52 MLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWL 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWL 60

Qy        112 GGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLV 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLV 120

Qy        172 AIIIGSPMWHVQRLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSV 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AIIIGSPMWHVQRLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSV 180

Qy        232 LYGKIGYELWIKKRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIV 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LYGKIGYELWIKKRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIV 240

Qy        292 HMMIEYSNFEKEYDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIV 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HMMIEYSNFEKEYDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIV 300

Qy        352 KETPSSARKHGSSGAMVMHRRAKLAARENPVEIKGEAFGGSNIDIKWCEQPEKKKRRSKV 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KETPSSARKHGSSGAMVMHRRAKLAARENPVEIKGEAFGGSNIDIKWCEQPEKKKRRSKV 360

Qy        412 ASCPL 416
              |||||
Db        361 ASCPL 365
```

SEQUENCE COMPARISON A

AC AAB02853;
XX
DT 22-AUG-2000 (first entry)
XX
DE Human G protein coupled receptor hRUP4 (V272K) protein SEQ ID NO:128.
XX
KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW identification; agonist; screening; therapeutic; pharmaceutical;
KW mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200022131-A2.
XX
PD 20-APR-2000.
XX
PF 13-OCT-1999; 99WO-US24065.
XX
PR 13-OCT-1998; 98US-0170496.
PR 12-NOV-1998; 98US-0108029.
PR 20-NOV-1998; 98US-0109213.
PR 27-NOV-1998; 98US-0110060.
PR 16-FEB-1999; 99US-0120416.
PR 26-FEB-1999; 99US-0121852.
PR 12-MAR-1999; 99US-0123944.
PR 12-MAR-1999; 99US-0123945.
PR 12-MAR-1999; 99US-0123946.
PR 12-MAR-1999; 99US-0123948.
PR 12-MAR-1999; 99US-0123949.
PR 12-MAR-1999; 99US-0123951.
PR 28-MAY-1999; 99US-0136436.
PR 28-MAY-1999; 99US-0136437.
PR 28-MAY-1999; 99US-0136439.
PR 28-MAY-1999; 99US-0137127.
PR 28-MAY-1999; 99US-0137131.
PR 28-MAY-1999; 99US-0137567.
PR 30-JUN-1999; 99US-0141448.
PR 27-AUG-1999; 99US-0151114.
PR 03-SEP-1999; 99US-0152524.
PR 29-SEP-1999; 99US-0156633.
PR 29-SEP-1999; 99US-0156555.
PR 29-SEP-1999; 99US-0156634.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
XX
DR WPI; 2000-317986/27.
DR N-PSDB; AAA46115.
XX
PT Non-endogenous, human G protein-coupled receptors for screening
PT receptor, inverse or partial agonists useful as therapeutic agents -

XX
PS Example 2; Page 164-166; 187pp; English.
XX
CC The present invention describes transmembrane receptors, preferably
CC human G protein coupled receptors (GPCR), for which the endogenous
CC ligand is unknown (orphan GPCR receptors). More specifically the present
CC invention relates to non-endogenous, constitutively activated versions
CC of a human GPCR. These non-endogenous human GPCRs can be useful for
CC the direct identification of candidate compounds as receptors agonists,
CC inverse agonists or partial agonists for use as pharmaceutical agents.
CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC the exemplification of the present invention.
XX
SQ Sequence 431 AA;

Query Match 81.4%; Score 1758; DB 21; Length 431;
Best Local Similarity 79.7%; Pred. No. 1.2e-187;
Matches 337; Conservative 37; Mismatches 39; Indels 10; Gaps 3;

```
Qy          4 NLTAEQLSALLRLHNLTRAQFIAHYGLRPLVLTPQLPARARLALLLVGMLIFALALFGNA 63
              |:| || | ||| ||||| |||| | ||||| ||:|| ||:|||:| |:||||||||||
Db          5 NITPEQFSRLLLRDHNLTREQFIALYRLRPLVYTPELPGRAKLALVLTGVLIFALALFGNA 64

Qy         64 LVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWLGGAFICKMVPFV 123
              || |||||||||||||||||||||||||| |||||||||||:|| |||||||||||||
Db         65 LVFYVVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLQNISDNWLGGAFICKMVPFV 124

Qy        124 QCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLVAIIIGSPMWHVQ 183
              | ||:|||:|||||||||||||||||| ||||:||||||||||||:|:||||||||
Db        125 QSTAVVTEMLTMTCIAVERHQGLVHPFKMKWQYTNRRAFTMLGVVWLVAVIVGSPMWHVQ 184

Qy        184 RLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSVLYGKIGYELWIK 243
              :|||||||||||||||||:||||||||||||| |||||::: :|| |||||||||
Db        185 QLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFILVILFLLPLMVMLILYSKIGYELWIK 244

Qy        244 KRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIVHMMIEYSNFEKE 303
              ||:|||||||||||||| |||||||| ||||||| ||||||||||:||||||||||||||
Db        245 KRVGDGSVLRTIHGKEMSKIARKKKRAKIMMVTVVALFAVCWAPFHVVHMMIEYSNFEKE 304

Qy        304 YDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIVKETPSSARKHGS 363
              ||:|||||||||||||| |||||||:|| |||||||| :|||||||| :| | |::||:
Db        305 YDDVTIKMIFAIVQIIGFSNSICNPIVYAFMNENFKKNVLSAVCYCIVNKTFSPAQRHGN 364

Qy        364 SGAMVMHRRAKLAARENPV-EIKGEAFGGSNIDIKWCEQPEKKKR--------RSKVA-S 413
              ||  :| ::|| : ||||| | |||||   ||::| ||| |:||:        ||::| :
Db        365 SGITMMRKKAKFSLRENPVEETKGEAFSDGNIEVKLCEQTEEKKKLKRHLALFRSELAEN 424

Qy        414 CPL 416
               ||
Db        425 SPL 427
```

SEQUENCE COMPARISON- B

neuropeptide Y/peptide YY receptor Y2 - human
N;Alternate names: neuropeptide y/peptide YY receptor type 2
C;Species: Homo sapiens (man)
C;Date: 01-Mar-1996 #sequence_revision 01-Mar-1996 #text_change 20-Apr-2000
C;Accession: I39187; I39163; G02301
R;Gerald, C.; Walker, M.W.; Vaysse, P.J.
J. Biol. Chem. 270, 26758-26761, 1995
A;Title: Expression cloning and pharmacological characterization of a human hippocampal neuropeptide Y/peptide YY Y2 receptor subtype.
A;Reference number: I39187; MUID:96070760; PMID:7592910
A;Accession: I39187
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-381 <GER>
A;Cross-references: EMBL:U36269; NID:g1063633; PIDN:AAC50281.1; PID:g1063634
R;Rose, P.M.; Fernandes, P.; Lynch, J.S.; Frazier, S.T.; Fisher, S.M.; Kodukula, K.; Kienzle, B.; Seethala, R.
J. Biol. Chem. 270, 22661-22664, 1995
A;Title: Cloning and functional expression of a cDNA encoding a human type 2 neuropeptide Y receptor.
A;Gene: GDB:NPY2R
A;Cross-references: GDB:4365607; OMIM:162642
A;Map position: 4q31-4q31
C;Superfamily: neurokinin 1 receptor
F;49-76/Domain: transmembrane #status predicted <TM1>
F;87-113/Domain: transmembrane #status predicted <TM2>
F;166-186/Domain: transmembrane #status predicted <TM4>
F;221-237/Domain: transmembrane #status predicted <TM5>
F;269-291/Domain: transmembrane #status predicted <TM6>
F;305-328/Domain: transmembrane #status predicted <TM7>
F;123-203/Disulfide bonds: #status predicted
F;342/Binding site: palmitate (Cys) (covalent) #status predicted
F;372/Binding site: carbohydrate (Asn) (covalent) #status predicted

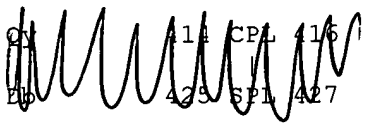
Query Match 20.5%; Score 442; DB 2; Length 381;
Best Local Similarity 30.7%; Pred. No. 3.7e-30;
Matches 96; Conservative 77; Mismatches 106; Indels 34; Gaps 9;

```
Qy         44 RLALLLVGMLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTML 103
              :: |:|    |  | : ||:||::|| : |:|||||| || :||::|||:   |:| |:
Db         50 QVVLILAYCSIILLGVIGNSLVIHVVIKFKSMRTVTNFFIANLAVADLLVNTLCLPFTLT 109

Qy        104 QNVSDTWLGGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFT 163
                :   |  |  :| :||: |  |:    :|:| ||::||: :|  : :: : : : :|
Db        110 YTLMGEWKMGPVLCHLVPYAQGLAVQVSTITLTVIALDRHRCIV--YHLESKISKRISFL 167

Qy        164 MLGVVWLVAIIIGSPM---WHVQRLEIKYDFLYEKEHICCLEEW---SSPVHQKIYTTFI 217
              ::|: | :: :: ||:        :||  ||    | : | |:|      ::  :|:
Db        168 IIGLAWGISALLASPLAIFREYSLIEIIPDF----EIVACTEKWPGEEKSIYGTVYSLSS 223

Qy        218 LVTLFLLPLLLLSVLYGKIGYELW--IKKRIGDGSVLRTIHGKEMFKIARKKKRAVIMMV 275
              |: |::||| ::|  | :|    |  :|  :  |:     |        :::::   |:|
Db        224 LLILYVLPLGIISFSYTRI----WSKLKNHVSPGAANDHYH--------QRRQKTTKMLV 271

Qy        276 TVVVLFAVCWAPFHIVHMMIEYSNFE---KEYDEVTIKMIFAIVQIIGFFNSICNPIIYA 332
               |||:||| | | |   : ::  :     |||     |:|| :  ||   ::  ||::|
```

```
Db        272 CVVVVFAVSWLPLHAFQLAVDIDSQVLDLKEY-----KLIFTVFHIIAMCSTFANPLLYG 326

Qy        333 LMNENFKKNFVSA 345   8/9=88% identical
                || |::| |:||
Db        327 WMNSNYRKAFLSA 339
```

SEQUENCE COMPARISON — C

```
AC   Q8BHH0; PRELIMINARY;      PRT;   365 AA
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical rhodopsin-like GPCR superfamily containing protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion, and Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK048439; BAC33337.1; -.
DR   EMBL; AK051723; BAC34735.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   365 AA;  41759 MW;  1EB7E5369632ED56 CRC64;

  Query Match             88.2%;  Score 1905;  DB 11;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-160;
  Matches  365;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps 0

Qy         52 MLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWL 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWL 60

Qy        112 GGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLV 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLV 120

Qy        172 AIIIGSPMWHVQRLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSV 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AIIIGSPMWHVQRLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSV 180

Qy        232 LYGKIGYELWIKKRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIV 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LYGKIGYELWIKKRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIV 240

Qy        292 HMMIEYSNFEKEYDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIV 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HMMIEYSNFEKEYDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIV 300

Qy        352 KETPSSARKHGSSGAMVMHRRAKLAARENPVEIKGEAFGGSNIDIKWCEQPEKKKRRSKV 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KETPSSARKHGSSGAMVMHRRAKLAARENPVEIKGEAFGGSNIDIKWCEQPEKKKRRSKV 360

Qy        412 ASCPL 416
              |||||
Db        361 ASCPL 365
```



SEQUENCE COMPARISON - D

```
AC   AAY71309;
XX
DT   02-NOV-2000  (first entry)
XX
DE   Human orphan G protein-coupled receptor hRUP4.
XX
KW   Human; orphan G protein-coupled receptor; GPCR; hRUP4; drug screening;
KW   transmembrane receptor; expressed sequence tag; EST; signal cascade.
XX
OS   Homo sapiens.
XX
PN   WO200031258-A2.
XX
PD   02-JUN-2000.
XX
PF   13-OCT-1999;   99WO-US23687.
XX
PR   20-NOV-1998;   98US-0109213.
PR   16-FEB-1999;   99US-0120416.
PR   26-FEB-1999;   99US-0121852.
PR   12-MAR-1999;   99US-0123946.
PR   12-MAR-1999;   99US-0123949.
PR   28-MAY-1999;   99US-0136436.
PR   28-MAY-1999;   99US-0136437.
PR   28-MAY-1999;   99US-0136439.
PR   28-MAY-1999;   99US-0136567.
PR   28-MAY-1999;   99US-0137127.
PR   28-MAY-1999;   99US-0137131.
PR   29-JUN-1999;   99US-0141448.
PR   29-SEP-1999;   99US-0156555.
PR   29-SEP-1999;   99US-0156633.
PR   29-SEP-1999;   99US-0156634.
PR   29-SEP-1999;   99US-0156653.
PR   01-OCT-1999;   99US-0157280.
PR   01-OCT-1999;   99US-0157281.
PR   01-OCT-1999;   99US-0157282.
PR   01-OCT-1999;   99US-0157293.
PR   01-OCT-1999;   99US-0157294.
PR   12-OCT-1999;   99US-0416760.
PR   12-OCT-1999;   99US-0417044.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R,  Dang HT,  Liaw CW,  Lin I;
XX
DR   WPI; 2000-400068/34.
```

DR N-PSDB; AAD01136.
XX
PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
PT for use in the identification of G protein-coupled receptor agonists -
XX
PS Claim 74; Page 89-91; 102pp; English.
XX
CC The present amino acid sequence is the hRUP4, an endogenous human
CC orphan G protein-coupled receptor (GPCR). The full length hRUP4 cDNA was
CC cloned by RT-PCR with human brain cDNA as template. The hRUP4 PCR
CC fragment obtained was an alternatively spliced form of the EST (expressed
CC sequence tag) clone AI307658. The orphan GPCR of the invention, like
CC all GPCRs has seven transmembrane alpha helices with an extracellular
CC N-terminus and an intracellular C-terminus. However, no endogenous
CC ligands has yet been identified for the proteins of the invention. The
CC orphan GPCRs may be used in the identification of their endogenous
CC ligands, and to screen potential GPCR agonists and antagonists for use as
CC pharmaceutical agents. The proteins may also be used in the study of
CC GPCR-mediated signalling cascades, and to elucidate their precise role in
CC normal and diseased human conditions. Nucleic acid encoding human orphan
CC GPCRs may be used for tissue localisation expression analysis to provide
CC information about their function in healthy and pathological states.
XX
SQ Sequence 431 AA;

Query Match 81.7%; Score 1764; DB 21; Length 431;
Best Local Similarity 79.9%; Pred. No. 2.5e-188;
Matches 338; Conservative 37; Mismatches 38; Indels 10; Gaps 3;

```
Qy          4 NLTAEQLSALLRLHNLTRAQFIAHYGLRPLVLTPQLPARARLALLLVGMLIFALALFGNA 63
              |:| || | ||| ||||| |||| | ||||| ||:|| ||:|||:| |:||||||||||
Db          5 NITPEQFSRLLLRDHNLTREQFIALYRLRPLVYTPELPGRAKLALVLTGVLIFALALFGNA 64

Qy         64 LVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWLGGAFICKMVPFV 123
              || |||||||||||||||||||||||||| ||||||||||:|| ||||||||||||||
Db         65 LVFYVVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLQNISDNWLGGAFICKMVPFV 124

Qy        124 QCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLVAIIIGSPMWHVQ 183
              | ||:|||:|||||||||||||||||||| ||||:||||||||||||:|:||||||||
Db        125 QSTAVVTEMLTMTCIAVERHQGLVHPFKMKWQYTNRRAFTMLGVVWLVAVIVGSPMWHVQ 184

Qy        184 RLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSVLYGKIGYELWIK 243
              :|||||||||||||||||:|||||||||||||| |||||||:: :|| |||||||||
Db        185 QLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFILVILFLLPLMVMLILYSKIGYELWIK 244

Qy        244 KRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIVHMMIEYSNFEKE 303
              ||:|||||||||||||| |||||||||||||||| |||||||||:||||||||||||||
Db        245 KRVGDGSVLRTIHGKEMSKIARKKKRAVIMMVTVVALFAVCWAPFHVVHMMIEYSNFEKE 304

Qy        304 YDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIVKETPSSARKHGS 363
              ||:|||||||||||||| |||||||:|| |||||||| :|||||||| :| | |::||:
Db        305 YDDVTIKMIFAIVQIIGFSNSICNPIVYAFMNENFKKNVLSAVCYCIVNKTFSPAQRHGN 364
```

```
Qy        364 SGAMVMHRRAKLAARENPV-EIKGEAFGGSNIDIKWCEQPEKKKR--------RSKVA-S 413
              ||  :| ::|| : ||||| | |||||   ||::| ||| |:||:        ||::| :
Db        365 SGITMMRKKAKFSLRENPVEETKGEAFSDGNIEVKLCEQTEEKKKLKRHLALFRSELAEN 424

Qy        414 CPL 416
               ||
Db        425 SPL 427
```

SEQUENCE COMPARISON - B

J. Biol. Chem. 270, 22661-22664, 1995
A;Title: Cloning and functional expression of a cDNA encoding a human type 2 neuropeptide Y receptor.
A;Reference number: I39163; MUID:96032678; PMID:7559383
A;Accession: I39163
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-133,'A',135-381 <ROS>
A;Cross-references: EMBL:U32500; NID:g1000750; PIDN:AAA93170.1; PID:g1000751
R;Yan, H.; Yang, J.; Marasco, J.; Yamaguchi, K.; Brenner, S.; Collins, F.; Karbon, W.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01019
A;Accession: G02301
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-171,'G',173,'R',175-201,'P',203-208,'A',210-381 <YAN>
A;Cross-references: EMBL:U42389; NID:g1314329; PIDN:AAB07760.1; PID:g1314330
C;Genetics:
A;Gene: GDB:NPY2R
A;Cross-references: GDB:4365607; OMIM:162642
A;Map position: 4q31-4q31
C;Superfamily: neurokinin 1 receptor
C;Keywords: appetite; G protein-coupled receptor; glycoprotein; lipoprotein; thiolester bond; transmembrane protein
F;49-76/Domain: transmembrane #status predicted <TM1>
F;87-113/Domain: transmembrane #status predicted <TM2>
F;166-186/Domain: transmembrane #status predicted <TM4>
F;221-237/Domain: transmembrane #status predicted <TM5>
F;269-291/Domain: transmembrane #status predicted <TM6>
F;305-328/Domain: transmembrane #status predicted <TM7>
F;123-203/Disulfide bonds: #status predicted
F;342/Binding site: palmitate (Cys) (covalent) #status predicted
F;372/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.5%; Score 442; DB 2; Length 381;
Best Local Similarity 30.7%; Pred. No. 3.7e-30;
Matches 96; Conservative 77; Mismatches 106; Indels 34; Gaps 9;

```
Qy          44 RLALLLVGMLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTML 103
               :: |:|    |  | : ||:||::|| : |:|||||| || :||::|||:   |:| |:
Db          50 QVVLILAYCSIILLGVIGNSLVIHVVIKFKSMRTVTNFFIANLAVADLLVNTLCLPFTLT 109

Qy         104 QNVSDTWLGGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFT 163
                 :   |  |  :| :||: |  |:    :|:| ||::||: :|  : :: : : : :|
Db         110 YTLMGEWKMGPVLCHLVPYAQGLAVQVSTITLTVIALDRHRCIV--YHLESKISKRISFL 167

Qy         164 MLGVVWLVAIIIGSPM---WHVQRLEIKYDFLYEKEHICCLEEW---SSPVHQKIYTTFI 217
               ::|: | :: :: ||:        :||  ||    | : | |:|      ::  :|:
Db         168 IIGLAWGISALLASPLAIFREYSLIEIIPDF----EIVACTEKWPGEEKSIYGTVYSLSS 223

Qy         218 LVTLFLLPLLLLSVLYGKIGYELW--IKKRIGDGSVLRTIHGKEMFKIARKKKRAVIMMV 275
               |: |::||| ::|  | :|    |  :|  :  |:     |        :::::   |:|
```

```
Db        224 LLILYVLPLGIISFSYTRI----WSKLKNHVSPGAANDHYH--------QRRQKTTKMLV 271

Qy        276 TVVVLFAVCWAPFHIVHMMIEYSNFE---KEYDEVTIKMIFAIVQIIGFFNSICNPIIYA 332
              |||:||| | | |   : ::  :     |||     |:|| :  ||   ::  ||::|
Db        272 CVVVVFAVSWLPLHAFQLAVDIDSQVLDLKEY-----KLIFTVFHIIAMCSTFANPLLYG 326

Qy        333 LMNENFKKNFVSA 345   8/9=88% identical
              || |::| |:||
Db        327 WMNSNYRKAFLSA 339
```

AC AAY94993;
XX
DT 19-JUN-2000 (first entry)
XX
DE Human secreted protein vc38_1, SEQ ID NO:26.
XX
KW Human; secreted protein; cancer; tumour; cardiovascular disorder;
KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
KW neurodegenerative disease; asthma; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200011015-A1.
XX
PD 02-MAR-2000.
XX
PF 24-AUG-1999; 99WO-US19351.
XX
PR 24-AUG-1998; 98US-0097638.
PR 24-AUG-1998; 98US-0097659.
PR 09-SEP-1998; 98US-0099618.
PR 28-SEP-1998; 98US-0102092.
PR 25-NOV-1998; 98US-0109978.
PR 23-DEC-1998; 98US-0113645.
PR 23-DEC-1998; 98US-0113646.
PR 23-AUG-1999; 99US-0379246.
XX
PA (ALPH-) ALPHAGENE INC.
XX
PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX
DR WPI; 2000-224657/19.
XX
PT New secreted or transmembrane proteins and polynucleotides encoding
PT them, useful for treating neurodegenerative disorders, autoimmune
PT diseases and cancer -
XX
PS Claim 35; Page 284-285; 357pp; English.
XX
CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),
CC and cDNA sequences encoding them (AAA23423-A23462). The secreted
CC proteins of the invention include those that are thought to be only
CC partially secreted, i.e., transmembrane proteins. The proteins of the
CC invention may exhibit one or more activities selected from the following:
CC cytokine activity; cell proliferation; differentiation; immune
CC modulation; haematopoiesis regulation; tissue growth activity;
CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
CC and thrombolytic activity; anti-inflammatory activity; and tumour
CC inhibition activity. The proteins may be administered to patients as
CC vaccines, and the nucleotides may be used as part of a gene therapy

```
CC   regime. Diseases or conditions that may be treated using the proteins or
CC   nucleotides of the invention include autoimmune diseases; genetic
CC   disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
CC   fungal and viral infections, especially HIV; multiple sclerosis;
CC   rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
CC   insulin dependent diabetes mellitus; and allergic reactions such as
CC   asthma and anaemia. They may also be used for treating wounds, burns,
CC   ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
CC   disease, Parkinson's disease, Huntington's disease and amyotrophic
CC   lateral sclerosis (ALS). Proteins with activin/inhibin activity may
CC   additionally be useful as contraceptives. Nucleic acid sequences of the
CC   invention may be used in chromosome mapping, and as a source of
CC   diagnostic primers and probes. The present sequence represents one of the
CC   40 proteins of the invention.
XX
SQ   Sequence   431 AA;

  Query Match             81.8%;  Score 1767;  DB 21;  Length 431;
  Best Local Similarity   80.1%;  Pred. No. 1.2e-188;
  Matches  339;  Conservative   36;  Mismatches   38;  Indels   10;  Gaps
3;

Qy          4 NLTAEQLSALLRLHNLTRAQFIAHYGLRPLVLTPQLPARARLALLLVGMLIFALALFGNA 63
              |:| || | ||| ||||| |||| | ||||| ||:|| ||:|||:| |:||||||||||
Db          5 NITPEQFSRLLRDHNLTREQFIALYRLRPLVYTPELPGRAKLALVLTGVLIFALALFGNA 64

Qy         64 LVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWLGGAFICKMVPFV 123
              || |||||||||||||||||||||||||| |||||||||||:|| |||||||||||||
Db         65 LVFYVVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLQNISDNWLGGAFICKMVPFV 124

Qy        124 QCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLVAIIIGSPMWHVQ 183
              | ||:|||||||||||||||||||||||| ||||:|||||||||||||:|:||||||||
Db        125 QSTAVVTEILTMTCIAVERHQGLVHPFKMKWQYTNRRAFTMLGVVWLVAVIVGSPMWHVQ 184

Qy        184 RLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSVLYGKIGYELWIK 243
              :|||||||||||||||||||:|||||||||||||| |||||||:: :|| |||||||||
Db        185 QLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFILVILFLLPLMVMLILYSKIGYELWIK 244

Qy        244 KRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIVHMMIEYSNFEKE 303
              ||:||||||||||||| |||||||||||||||| |||||||||:||||||||||||||
Db        245 KRVGDGSVLRTIHGKEMSKIARKKKRAVIMMVTVVALFAVCWAPFHVVHMMIEYSNFEKE 304

Qy        304 YDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIVKETPSSARKHGS 363
              ||:|||||||||||||| |||||||:|| |||||||| :|||||||| :| | |::||:
Db        305 YDDVTIKMIFAIVQIIGFSNSICNPIVYAFMNENFKKNVLSAVCYCIVNKTFSPAQRHGN 364

Qy        364 SGAMVMHRRAKLAARENPV-EIKGEAFGGSNIDIKWCEQPEKKKR--------RSKVA-S 413
              ||  :| ::|| : ||||| | |||||   ||::| ||| |:||:        ||::| :
Db        365 SGITMMRKKAKFSLRENPVEETKGEAFSDGNIEVKLCEQTEEKKKLKRHLALFRSELAEN 424

Qy        414 CPL 416
               ||
Db        425 SPL 427
```

```
AC   Q9Z2D5;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN   NPY2R.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99017376; PubMed=9802390;
RA   Sharma P.S., Holmberg S.K., Eriksson H., Beck-Sickinger A.G.,
RA   Grundemar L., Larhammar D.;
RT   "Cloning and functional expression of the guinea pig neuropeptide Y
RT   Y2 receptor.";
RL   Regul. Pept. 75:23-28(1998).
CC   -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       HIGHEST TO TACHYKININS RECEPTORS.
CC   ----------------------------------------------------------------------
-
CC   This SWISS-PROT entry is copyright. It is produced through a
collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation
-
CC   the European Bioinformatics Institute.  There are no  restrictions on
its
CC   use  by  non-profit  institutions as long  as its content  is  in  no
way
CC   modified and this statement is not removed.  Usage  by  and for
commercial
CC   entities requires a license agreement (See http://www.isb-
sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
-
DR   EMBL; AF072821; AAD13143.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Lipoprotein; Palmitate.
```

```
FT   DOMAIN        1     50       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     51     73       1 (POTENTIAL).
FT   DOMAIN       74     83       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     84    105       2 (POTENTIAL).
FT   DOMAIN      106    125       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    126    147       3 (POTENTIAL).
FT   DOMAIN      148    167       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    168    188       4 (POTENTIAL).
FT   DOMAIN      189    215       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    216    241       5 (POTENTIAL).
FT   DOMAIN      242    269       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    270    292       6 (POTENTIAL).
FT   DOMAIN      293    305       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    306    329       7 (POTENTIAL).
FT   DOMAIN      330    381       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     11     11       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    124    204       BY SIMILARITY.
FT   LIPID       343    343       PALMITATE (POTENTIAL).
SQ   SEQUENCE   381 AA;  42342 MW;  D806B859A43ECACE CRC64;

  Query Match             20.9%;  Score 452;  DB 1;  Length 381;
  Best Local Similarity   31.6%;  Pred. No. 4.5e-25;
  Matches   99;  Conservative   74;  Mismatches  106;  Indels   34;  Gaps
9;

Qy          44 RLALLLVGMLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTML 103
               |: |:|    |  | : ||:||::|| : |:|||||| || :||::|||:   |:| |:
Db          51 RVVLILAYCSIILLGVVGNSLVIHVVIKFKSMRTVTNFFIANLAVADLLVNTLCLPFTLT 110

Qy         104 QNVSDTWLGGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFT 163
                 :   |  |  :| :||: |  |:    :|:| ||::||: :|  : :  : : | :|
Db         111 YTLMGEWKMGPVLCHLVPYAQGLAVQVSTVTLTVIALDRHRCIV--YHLDSKISKQNSFL 168

Qy         164 MLGVVWLVAIIIGSPM---WHVQRLEIKYDFLYEKEHICCLEEW---SSPVHQKIYTTFI 217
               ::|: | :: :: ||:        :||  ||    | : | |:|      ::  :|:
Db         169 IIGLAWGISALLASPLAIFREYSLIEIIPDF----EIVACTEKWPGEEKSIYGTVYSLSS 224

Qy         218 LVTLFLLPLLLLSVLYGKIGYELW--IKKRIGDGSVLRTIHGKEMFKIARKKKRAVIMMV 275
               |: |::||| ::|| | :|    |  :|  :  |:     |        :::::   |:|
Db         225 LLILYVLPLGIISVSYVRI----WSKLKNHVSPGAANDHYH--------QRRQKTTKMLV 272

Qy         276 TVVVLFAVCWAPFHIVHMMIEYSNFE---KEYDEVTIKMIFAIVQIIGFFNSICNPIIYA 332
                |||:||| | | |   : ::  :     |||     |:|| :  ||   ::  ||::|
Db         273 FVVVVFAVSWLPLHAFQLAVDIDSQVLDLKEY-----KLIFTVFHIIAMCSTFANPLLYG 327

Qy         333 LMNENFKKNFVSA 345
                || |::| |:||
Db         328 WMNSNYRKAFLSA 340

ID   NY2R_BOVIN     STANDARD;      PRT;   384 AA.
AC   P79113;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
```

```
DE   Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN   NPY2R.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ammar D.A., Kolakowski L.F. Jr., Eadie D.M., Wong D.J., Ma Y.Y.,
RA   Yang-Feng T.L., Thompson D.A.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       HIGHEST TO TACHYKININS RECEPTORS.
CC   ---------------------------------------------------------------------
-
CC   This SWISS-PROT entry is copyright. It is produced through a
collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
-
CC   the European Bioinformatics Institute. There are no restrictions on
its
CC   use by non-profit institutions as long as its content is in no
way
CC   modified and this statement is not removed. Usage by and for
commercial
CC   entities requires a license agreement (See http://www.isb-
sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
-
DR   EMBL; U50144; AAB40600.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     53     75       1 (POTENTIAL).
FT   DOMAIN       76     85       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     86    107       2 (POTENTIAL).
FT   DOMAIN      108    127       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    128    149       3 (POTENTIAL).
FT   DOMAIN      150    169       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    170    190       4 (POTENTIAL).
FT   DOMAIN      191    217       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    218    243       5 (POTENTIAL).
FT   DOMAIN      244    271       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    272    294       6 (POTENTIAL).
FT   DOMAIN      295    307       EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM    308    331       7 (POTENTIAL).
FT   DOMAIN      332    384       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     13     13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    126    206       BY SIMILARITY.
FT   LIPID       345    345       PALMITATE (POTENTIAL).
SQ   SEQUENCE   384 AA;  42943 MW;  468D19CBA8F29681 CRC64;

  Query Match             20.8%;  Score 449;  DB 1;  Length 384;
  Best Local Similarity   31.6%;  Pred. No. 7.4e-25;
  Matches   99;  Conservative   74;  Mismatches  106;  Indels   34;  Gaps
9;

Qy         44 RLALLLVGMLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTML 103
              :: |:|    |  | : ||:||::|| : |:|||||| || :||::|||:   |:| |:
Db         53 QVVLILAYCSIILLGVIGNSLVIHVVIKFKSMRTVTNFFIANLAVADLLVNTLCLPFTLT 112

Qy        104 QNVSDTWLGGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFT 163
                :   |  |  :| :||: |  |:    :|:| ||::||: :|  : :: : : | :|
Db        113 YTLMGEWKMGPVLCHLVPYAQGLAVQVSTITLTVIALDRHRCIV--YHLESKISKQISFL 170

Qy        164 MLGVVWLVAIIIGSPM---WHVQRLEIKYDFLYEKEHICCLEEW---SSPVHQKIYTTFI 217
              ::|: | |: :: ||:        :||  ||    | : | |:|      ::  ||:
Db        171 IIGLAWGVSALLASPLAIFREYSLIEIIPDF----EIVACTEKWPGEEKGIYGTIYSLSS 226

Qy        218 LVTLFLLPLLLLSVLYGKIGYELW--IKKRIGDGSVLRTIHGKEMFKIARKKKRAVIMMV 275
              |: |::||| ::|  | :|    |  :|  :  |:     |        :::::   |:|
Db        227 LLILYVLPLGIISFSYTRI----WSKLKNHVSPGAAHDHYH--------QRRQKTTKMLV 274

Qy        276 TVVVLFAVCWAPFHIVHMMIEYSNFE---KEYDEVTIKMIFAIVQIIGFFNSICNPIIYA 332
               |||:||| | | |   : ::  :     |||     |:|| :  ||   ::  ||::|
Db        275 CVVVVFAVSWLPLHAFQLAVDIDSHVLDLKEY-----KLIFTVFHIIAMCSTFANPLLYG 329

Qy        333 LMNENFKKNFVSA 345
               || |::| |:||
Db        330 WMNSNYRKAFLSA 342


neuropeptide Y/peptide YY receptor Y2 - human
N;Alternate names: neuropeptide y/peptide YY receptor type 2
C;Species: Homo sapiens (man)
C;Date: 01-Mar-1996 #sequence_revision 01-Mar-1996 #text_change 20-Apr-2000
C;Accession: I39187; I39163; G02301
R;Gerald, C.; Walker, M.W.; Vaysse, P.J.
J. Biol. Chem. 270, 26758-26761, 1995
A;Title: Expression cloning and pharmacological characterization of a human
hippocampal neuropeptide Y/peptide YY Y2 receptor subtype.
A;Reference number: I39187; MUID:96070760; PMID:7592910
A;Accession: I39187
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-381 <GER>
A;Cross-references: EMBL:U36269; NID:g1063633; PIDN:AAC50281.1; PID:g1063634
R;Rose, P.M.; Fernandes, P.; Lynch, J.S.; Frazier, S.T.; Fisher, S.M.;
Kodukula, K.; Kienzle, B.; Seethala, R.
```

; Patent No. 5989834
; GENERAL INFORMATION:
; APPLICANT: Synaptic Pharmaceutical Corporation
; TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE
; TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,355A
; FILING DATE: No. 5989834ember 26, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPW/MAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-687-355A-6

Query Match 21.1%; Score 454.5; DB 2; Length 381;
Best Local Similarity 30.1%; Pred. No. 3.7e-33;
Matches 104; Conservative 75; Mismatches 119; Indels 47; Gaps 10;

```
Qy          25 IAHYGLRPLVLTPQLP-------------ARARLALLLVGMLIFALALFGNALVVYVVTR 71
               :  ||  |     :||                :: |:|    |  | : ||:||::|| :
Db          18 VEFYGSGPTTPRGELPPDPEPELIDSTKLVEVQVVLILAYCSIILLGVVGNSLVIHVVIK 77

Qy          72 SKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWLGGAFICKMVPFVQCTAIVTE 131
                |:|||||| || :||::||||:   |:| |:   :   |  |  :| :||: |  |:
Db          78 FKSMRTVTNFFIANLAVADLLVNTLCLPFTLTYTLMGEWKMGPVLCHLVPYAQGLAVQVS 137

Qy         132 ILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLVAIIIGSPM---WHVQRLEIK 188
                :|:| ||::||: :|  : :: : : | :| ::|: | |: :: ||:        :||
Db         138 TITLTVIALDRHRCIV--YHLESKISKQISFLIIGLAWGVSALLASPLAIFREYSLIEII 195
```

```
Qy        189 YDFLYEKEHICCLEEW---SSPVHQKIYTTFILVTLFLLPLLLLSVLYGKIGYELW--IK 243
              ||    | : | |:|      |:  :|:   |: |::||| ::|  | :|    |  :|
Db        196 PDF----EIVACTEKWPGEEKSVYGTVYSLSTLLILYVLPLGIISFSYTRI----WSKLK 247

Qy        244 KRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIVHMMIEYSNFE-- 301
                :  |:     |        ::: :   |:| |||:||| | | |   : ::  :
Db        248 NHVSPGAASDHYH--------QRRHKMTKMLVCVVVVFAVSWLPLHAFQLAVDIDSHVLD 299

Qy        302 -KEYDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSA 345
              |||     |:|| :  ||   ::  ||::|  || |::| |:||
Db        300 LKEY-----KLIFTVFHIIAMCSTFANPLLYGWMNSNYRKAFLSA 339
```

```
; Sequence 2, Application US/09899532
; Patent No. US20020048791A1
; GENERAL INFORMATION:
;  APPLICANT: Bloomquist, Brian T.
;  APPLICANT:  Zhelnin, Leonid
;  TITLE OF INVENTION: Human Neuropeptide Y-Like G
;  TITLE OF INVENTION:  Protein-Coupled Receptor
;  FILE REFERENCE: 02973.00040
;  CURRENT APPLICATION NUMBER: US/09/899,532
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: US 60/216,523
;  PRIOR FILING DATE: 2000-07-06
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 431
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-899-532-2

  Query Match             81.8%;  Score 1767;  DB 9;  Length 431;
  Best Local Similarity   80.1%;  Pred. No. 7.3e-156;
  Matches  339;  Conservative   36;  Mismatches   38;  Indels   10;  Gaps
3;

Qy          4 NLTAEQLSALLRLHNLTRAQFIAHYGLRPLVLTPQLPARARLALLLVGMLIFALALFGNA 63
              |:| || | ||| ||||| |||| | ||||| ||:|| ||:|||:| |:||||||||||
Db          5 NITPEQFSRLLRDHNLTREQFIALYRLRPLVYTPELPGRAKLALVLTGVLIFALALFGNA 64

Qy         64 LVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWLGGAFICKMVPFV 123
              || |||||||||||||||||||||||||| |||||||||||:|| |||||||||||||
Db         65 LVFYVVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLQNISDNWLGGAFICKMVPFV 124

Qy        124 QCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLVAIIIGSPMWHVQ 183
              | ||:||||||||||||||||||||||| ||||:|||||||||||||:|:||||||||
Db        125 QSTAVVTEILTMTCIAVERHQGLVHPFKMKWQYTNRRAFTMLGVVWLVAVIVGSPMWHVQ 184

Qy        184 RLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSVLYGKIGYELWIK 243
              :|||||||||||||||||||:||||||||||||| |||||||:: :|| |||||||||
Db        185 QLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFILVILFLLPLMVMLILYSKIGYELWIK 244

Qy        244 KRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIVHMMIEYSNFEKE 303
              ||:|||||||||||||| |||||||||||||||| ||||||||||:||||||||||||||
Db        245 KRVGDGSVLRTIHGKEMSKIARKKKRAVIMMVTVVALFAVCWAPFHVVHMMIEYSNFEKE 304

Qy        304 YDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIVKETPSSARKHGS 363
              ||:|||||||||||||| |||||||:|| |||||||| :|||||||| :| | |::||:
Db        305 YDDVTIKMIFAIVQIIGFSNSICNPIVYAFMNENFKKNVLSAVCYCIVNKTFSPAQRHGN 364

Qy        364 SGAMVMHRRAKLAARENPV-EIKGEAFGGSNIDIKWCEQPEKKKR--------RSKVA-S 413
              ||  :| ::|| : ||||| | |||||   ||::| ||| |:||:        ||::| :
Db        365 SGITMMRKKAKFSLRENPVEETKGEAFSDGNIEVKLCEQTEEKKKLKRHLALFRSELAEN 424
```

```
RESULT 5
US-10-272-983-38
; Sequence 38, Application US/10272983
; Publication No. US20030148450A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Ruoping
;  APPLICANT:  Dang, Huong T.
;  APPLICANT:  Liaw, Chen W.
;  APPLICANT:  Lin, I-Lin
;  TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
;  FILE REFERENCE: AREN0050
;  CURRENT APPLICATION NUMBER: US/10/272,983
;  CURRENT FILING DATE:  2002-10-17
;  PRIOR APPLICATION NUMBER: US/09/417,044
;  PRIOR FILING DATE: 1999-10-12
;  PRIOR APPLICATION NUMBER: 60/109,213
;  PRIOR FILING DATE: 1998-11-20
;  PRIOR APPLICATION NUMBER: 60/120,416
;  PRIOR FILING DATE: 1999-02-16
;  PRIOR APPLICATION NUMBER: 60/121,851
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 60/123,946
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123,949
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/136,436
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,437
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,439
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,567
;  PRIOR FILING DATE: 1999-05-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
;   LENGTH: 431
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-272-983-38

  Query Match             81.7%;  Score 1764;  DB 12;  Length 431;
  Best Local Similarity   79.9%;  Pred. No. 1.4e-155;
  Matches  338;  Conservative   37;  Mismatches   38;  Indels   10;  Gaps
3;

Qy          4 NLTAEQLSALLRLHNLTRAQFIAHYGLRPLVLTPQLPARARLALLLVGMLIFALALFGNA 63
              |:| || | ||| ||||| |||| | ||||| ||:|| ||:|||:| |:||||||||||
Db          5 NITPEQFSRLLRDHNLTREQFIALYRLRPLVYTPELPGRAKLALVLTGVLIFALALFGNA 64

Qy         64 LVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWLGGAFICKMVPFV 123
              || |||||||||||||||||||||||||| ||||||||||:|| ||||||||||||||
```

```
Db        65 LVFYVVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLQNISDNWLGGAFICKMVPFV 124

Qy       124 QCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLVAIIIGSPMWHVQ 183
            | ||:|||:||||||||||||||||||||| ||||:||||||||||||:|:||||||||
Db       125 QSTAVVTEMLTMTCIAVERHQGLVHPFKMKWQYTNRRAFTMLGVVWLVAVIVGSPMWHVQ 184

Qy       184 RLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSVLYGKIGYELWIK 243
            :|||||||||||||||||||:|||||||||||||| |||||||:: :|| ||||||||||
Db       185 QLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFILVILFLLPLMVMLILYSKIGYELWIK 244

Qy       244 KRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIVHMMIEYSNFEKE 303
            ||:|||||||||||||| |||||||||||||||| |||||||||:||||||||||||||
Db       245 KRVGDGSVLRTIHGKEMSKIARKKKRAVIMMVTVVALFAVCWAPFHVVHMMIEYSNFEKE 304

Qy       304 YDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIVKETPSSARKHGS 363
            ||:|||||||||||||| |||||||:|| |||||||| :|||||||| :| | |::||:
Db       305 YDDVTIKMIFAIVQIIGFSNSICNPIVYAFMNENFKKNVLSAVCYCIVNKTFSPAQRHGN 364

Qy       364 SGAMVMHRRAKLAARENPV-EIKGEAFGGSNIDIKWCEQPEKKKR--------RSKVA-S 413
            ||  :| ::|| : ||||| | |||||   ||::| ||| |:||:        ||::| :
Db       365 SGITMMRKKAKFSLRENPVEETKGEAFSDGNIEVKLCEQTEEKKKLKRHLALFRSELAEN 424

Qy       414 CPL 416
             ||
Db       425 SPL 427
```

```
LOCUS       AK048439               2864 bp    mRNA    linear   HTC 05-DEC-
2002
DEFINITION  Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
            library, clone:C130060K24 product:hypothetical Rhodopsin-like GPCR
            superfamily containing protein, full insert sequence.
ACCESSION   AK048439
VERSION     AK048439.1  GI:26339331
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
```

MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2864)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,

Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source 1. .2864
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C130060K24"
/db_xref="taxon:10090"
/clone="C130060K24"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"

CDS 64. .1161
/note="unnamed protein product; hypothetical Rhodopsin-like GPCR superfamily containing protein (InterPro|IPR000276, evidence: InterPro) putative"
/codon_start=1
/protein_id="BAC33337.1"
/db_xref="GI:26339332"

/translation="MLIFALALFGNALVVYVVTRSKAMRTVTNIFICSLALSDLLIVF

FCIPVTMLQNVSDTWLGGAFICKMVPFVQCTAIVTEILTMTCIAVERHQGLVHPFKMK

RQYTNQRAFTMLGVVWLVAIIIGSPMWHVQRLEIKYDFLYEKEHICCLEEWSSPVHQK

IYTTFILVTLFLLPLLLLSVLYGKIGYELWIKKRIGDGSVLRTIHGKEMFKIARKKKR

AVIMMVTVVVLFAVCWAPFHIVHMMIEYSNFEKEYDEVTIKMIFAIVQIIGFFNSICN

PIIYALMNENFKKNFVSAVCYCIVKETPSSARKHGSSGAMVMHRRAKLAARENPVEIK
GEAFGGSNIDIKWCEQPEKKKRRSKVASCPL"
BASE COUNT 839 a 609 c 602 g 814 t
ORIGIN

Query Match 94.9%; Score 1681; DB 11; Length 2864;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        92 GGCCACTGGTGCTCACCCCGCAGCTTCCCGCGCGCGCCAGGCTGGCCCTCCTGCTGGTCG 151
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2 GGCCACTGGTGCTCACCCCGCAGCTTCCCGCGCGCGCCAGGCTGGCCCTCCTGCTGGTCG 61

Qy       152 GCATGCTCATCTTTGCCCTGGCGCTCTTCGGCAACGCCCTGGTAGTCTATGTGGTGACCC 211
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        62 GCATGCTCATCTTTGCCCTGGCGCTCTTCGGCAACGCCCTGGTAGTCTATGTGGTGACCC 121

Qy       212 GCAGCAAGGCCATGCGCACCGTCACCAACATCTTCATCTGCTCCCTGGCACTCAGCGACC 271
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       122 GCAGCAAGGCCATGCGCACCGTCACCAACATCTTCATCTGCTCCCTGGCACTCAGCGACC 181

Qy       272 TGCTCATCGTCTTCTTCTGCATCCCGGTCACCATGCTCCAGAACGTCTCGGACACCTGGC 331
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       182 TGCTCATCGTCTTCTTCTGCATCCCGGTCACCATGCTCCAGAACGTCTCGGACACCTGGC 241

Qy       332 TGGGGGGTGCCTTCATTTGCAAAATGGTCCCATTTGTCCAGTGCACTGCCATTGTGACAG 391
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       242 TGGGGGGTGCCTTCATTTGCAAAATGGTCCCATTTGTCCAGTGCACTGCCATTGTGACAG 301

Qy       392 AAATCCTTACTATGACCTGCATTGCTGTGGAAAGGCACCAGGGACTTGTCCATCCTTTTA 451
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       302 AAATCCTTACTATGACCTGCATTGCTGTGGAAAGGCACCAGGGACTTGTCCATCCTTTTA 361

Qy       452 AAATGAAGCGGCAGTACACCAATCAAAGAGCTTTCACAATGCTAGGTGTGGTGTGGCTGG 511
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       362 AAATGAAGCGGCAGTACACCAATCAAAGAGCTTTCACAATGCTAGGTGTGGTGTGGCTGG 421

Qy       512 TGGCCATCATCATAGGATCACCCATGTGGCATGTGCAGCGACTTGAGATTAAGTATGACT 571
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       422 TGGCCATCATCATAGGATCACCCATGTGGCATGTGCAGCGACTTGAGATTAAGTATGACT 481

Qy       572 TCCTATATGAAAAAGAACACATCTGCTGCCTGGAAGAGTGGAGCAGCCCCGTGCACCAGA 631
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       482 TCCTATATGAAAAAGAACACATCTGCTGCCTGGAAGAGTGGAGCAGCCCCGTGCACCAGA 541

Qy       632 AGATCTACACCACCTTCATCCTTGTCACCCTCTTCCTGCTACCACTGTTGCTGCTCTCTG 691
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       542 AGATCTACACCACCTTCATCCTTGTCACCCTCTTCCTGCTACCACTGTTGCTGCTCTCTG 601

Qy       692 TCCTCTACGGGAAAATCGGTTATGAGCTTTGGATCAAGAAAAGAATCGGGGATGGCTCAG 751
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 TCCTCTACGGGAAAATCGGTTATGAGCTTTGGATCAAGAAAAGAATCGGGGATGGCTCAG 661

Qy        752 TGCTCCGAACTATTCATGGAAAAGAAATGTTCAAAATAGCCAGAAAGAAGAAGCGAGCTG 811
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TGCTCCGAACTATTCATGGAAAAGAAATGTTCAAAATAGCCAGAAAGAAGAAGCGAGCTG 721

Qy        812 TGATCATGATGGTGACAGTCGTGGTTCTCTTTGCTGTGTGCTGGGCACCTTTCCACATCG 871
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 TGATCATGATGGTGACAGTCGTGGTTCTCTTTGCTGTGTGCTGGGCACCTTTCCACATCG 781

Qy        872 TTCACATGATGATTGAATACAGTAATTTTGAAAAGGAATATGATGAAGTCACAATCAAGA 931
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TTCACATGATGATTGAATACAGTAATTTTGAAAAGGAATATGATGAAGTCACAATCAAGA 841

Qy        932 TGATTTTTGCTATAGTGCAAATAATTGGATTTTTCAACTCCATCTGTAATCCCATTATTT 991
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 TGATTTTTGCTATAGTGCAAATAATTGGATTTTTCAACTCCATCTGTAATCCCATTATTT 901

Qy        992 ATGCACTTATGAATGAAAACTTCAAAAAAAAACTTTGTGTCTGCCGTTTGCTATTGCATTG
1051
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 ATGCACTTATGAATGAAAACTTCAAAAAAAAACTTTGTGTCTGCCGTTTGCTATTGCATTG 961

Qy       1052 TGAAGGAAACACCTTCTTCAGCACGGAAGCATGGAAGTTCAGGAGCTATGGTGATGCACA
1111
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TGAAGGAAACACCTTCTTCAGCACGGAAGCATGGAAGTTCAGGAGCTATGGTGATGCACA
1021

Qy       1112 GGAGGGCAAAGTTAGCTGCAAGAGAGAATCCTGTAGAGATCAAAGGAGAAGCATTTGGGG
1171
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 GGAGGGCAAAGTTAGCTGCAAGAGAGAATCCTGTAGAGATCAAAGGAGAAGCATTTGGGG
1081

Qy       1172 GCAGCAACATCGATATCAAGTGGTGTGAACAGCCAGAAAAGAAGAAGAGGAGATCAAAAG
1231
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 GCAGCAACATCGATATCAAGTGGTGTGAACAGCCAGAAAAGAAGAAGAGGAGATCAAAAG
1141

Qy       1232 TGGCATCTTGTCCTCTTTAGTTCCGAATTTCTGAGAGCTCTGCTGTAGACGTGAACACTG
1291
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TGGCATCTTGTCCTCTTTAGTTCCGAATTTCTGAGAGCTCTGCTGTAGACGTGAACACTG
1201

Qy       1292 TACCAATGTCTTCAGAATGAGTATCTGTCATACTGTAATCGAAAGAAAATGATTTTGAGA
1351
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TACCAATGTCTTCAGAATGAGTATCTGTCATACTGTAATCGAAAGAAAATGATTTTGAGA
1261
```

```
Qy       1352 AAAAGCCAGAGAGCTTTCATATTAAAAATGTTGACAAACACTCAGAAGGCAGGGACAGGG
1411

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAAAGCCAGAGAGCTTTCATATTAAAAATGTTGACAAACACTCAGAAGGCAGGGACAGGG
1321

Qy       1412 GATTCAAGAGTTTAAAGTCATCCTTAGCTGCACGATAAGTTTGAGGATAACCTGGGCTAC
1471

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 GATTCAAGAGTTTAAAGTCATCCTTAGCTGCACGATAAGTTTGAGGATAACCTGGGCTAC
1381

Qy       1472 AAGAGACCCTGTCTCAAGAAGCCATAATAATTAAAACAACCATCCTTAACTAATGATAAT
1531

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 AAGAGACCCTGTCTCAAGAAGCCATAATAATTAAAACAACCATCCTTAACTAATGATAAT
1441

Qy       1532 GACAAAGTATTTTTCCATTGAAAATACATGTAAGCTGCAATTTTGAAAAATTATTGAACC
1591

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 GACAAAGTATTTTTCCATTGAAAATACATGTAAGCTGCAATTTTGAAAAATTATTGAACC
1501

Qy       1592 ACCCTTGTGATTAATAGATGAAGTTTAAAAAAATTTAAATGTGTTTTTATTGTATGTATA
1651

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 ACCCTTGTGATTAATAGATGAAGTTTAAAAAAATTTAAATGTGTTTTTATTGTATGTATA
1561

Qy       1652 TGGTTGGTTTACCTGTGTATATGTCTTTCAGTAACTTGTATAAAACTCAATGATCTCAGC
1711

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 TGGTTGGTTTACCTGTGTATATGTCTTTCAGTAACTTGTATAAAACTCAATGATCTCAGC
1621

Qy       1712 TAGTAACTTTCTTCTGTGTGGTCAATGTGATATGATTTCCTATATATTGCTAAATTGAAT
1771

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 TAGTAACTTTCTTCTGTGTGGTCAATGTGATATGATTTCCTATATATTGCTAAATTGAAT
1681

Qy       1772 G 1772
                |
Db       1682 G 1682
```

```
RESULT 5
US-10-272-983-38
; Sequence 38, Application US/10272983
; Publication No. US20030148450A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Ruoping
;  APPLICANT:  Dang, Huong T.
;  APPLICANT:  Liaw, Chen W.
;  APPLICANT:  Lin, I-Lin
;  TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
;  FILE REFERENCE: AREN0050
;  CURRENT APPLICATION NUMBER: US/10/272,983
;  CURRENT FILING DATE:  2002-10-17
;  PRIOR APPLICATION NUMBER: US/09/417,044
;  PRIOR FILING DATE: 1999-10-12
;  PRIOR APPLICATION NUMBER: 60/109,213
;  PRIOR FILING DATE: 1998-11-20
;  PRIOR APPLICATION NUMBER: 60/120,416
;  PRIOR FILING DATE: 1999-02-16
;  PRIOR APPLICATION NUMBER: 60/121,851
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 60/123,946
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123,949
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/136,436
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,437
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,439
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,567
;  PRIOR FILING DATE: 1999-05-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
;   LENGTH: 431
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-272-983-38

  Query Match             81.7%;  Score 1764;  DB 12;  Length 431;
  Best Local Similarity   79.9%;  Pred. No. 1.4e-155;
  Matches  338;  Conservative   37;  Mismatches   38;  Indels   10;  Gaps
3;

Qy          4 NLTAEQLSALLRLHNLTRAQFIAHYGLRPLVLTPQLPARARLALLLVGMLIFALALFGNA 63
              |:| || | ||| ||||| |||| | ||||| ||:|| ||:|||:| |:||||||||||
Db          5 NITPEQFSRLLRDHNLTREQFIALYRLRPLVYTPELPGRAKLALVLTGVLIFALALFGNA 64

Qy         64 LVVYVVTRSKAMRTVTNIFICSLALSDLLIVFFCIPVTMLQNVSDTWLGGAFICKMVPFV 123
              || ||||||||||||||||||||||||| ||||||||||:|| ||||||||||||||
```

```
Db        65 LVFYVVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLQNISDNWLGGAFICKMVPFV 124

Qy       124 QCTAIVTEILTMTCIAVERHQGLVHPFKMKRQYTNQRAFTMLGVVWLVAIIIGSPMWHVQ 183
             | ||:|||:|||||||||||||||||||| ||||:||||||||||||:|:||||||||
Db       125 QSTAVVTEMLTMTCIAVERHQGLVHPFKMKWQYTNRRAFTMLGVVWLVAVIVGSPMWHVQ 184

Qy       184 RLEIKYDFLYEKEHICCLEEWSSPVHQKIYTTFILVTLFLLPLLLLSVLYGKIGYELWIK 243
             :|||||||||||||||||||:|||||||||||||| |||||||:: :|| |||||||||
Db       185 QLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFILVILFLLPLMVMLILYSKIGYELWIK 244

Qy       244 KRIGDGSVLRTIHGKEMFKIARKKKRAVIMMVTVVVLFAVCWAPFHIVHMMIEYSNFEKE 303
             ||:|||||||||||||| |||||||||||||||| ||||||||||:||||||||||||||
Db       245 KRVGDGSVLRTIHGKEMSKIARKKKRAVIMMVTVVALFAVCWAPFHVVHMMIEYSNFEKE 304

Qy       304 YDEVTIKMIFAIVQIIGFFNSICNPIIYALMNENFKKNFVSAVCYCIVKETPSSARKHGS 363
             ||:|||||||||||||| |||||||:|| |||||||| :|||||||| :| | |::||:
Db       305 YDDVTIKMIFAIVQIIGFSNSICNPIVYAFMNENFKKNVLSAVCYCIVNKTFSPAQRHGN 364

Qy       364 SGAMVMHRRAKLAARENPV-EIKGEAFGGSNIDIKWCEQPEKKKR--------RSKVA-S 413
             ||  :| ::|| : ||||| | |||||   ||::| ||| |:||:        ||::| :
Db       365 SGITMMRKKAKFSLRENPVEETKGEAFSDGNIEVKLCEQTEEKKKLKRHLALFRSELAEN 424

Qy       414 CPL 416
              ||
Db       425 SPL 427
```